

比喩解釈を目的とする World Wide Web を利用した属性値の適合性判定

梶井文人 †, ‡
三重大学工学部 †
masui@ai.info.mie-u.ac.jp †

福本淳一 ††
立命館大学情報理工学部 ††
fukumoto@cs.ritsumeai.ac.jp ††

荒木 健治 †
北海道大学大学院情報科学研究科 †
araki@media.eng.hokudai.ac.jp †

1 はじめに

比喩とは、ある概念を他の概念によって説明または強調する修辭的手法の一つであり、認知科学、言語学、心理学、人工知能など、様々な分野で研究対象として取り上げられている [1]。自然言語処理の分野においても例外ではなく、処理対象を一般的な文書や会話などへ拡大し、柔軟な処理を行うためには、比喩表現はしばしば問題となり、比喩表現を検出するだけでも有益である [2, 3, 4]。

比喩解釈の過程には、概念の属性値に注目するものと概念の構造に注目するものがあるが、本論文では、概念の属性値に基づく比喩 (属性比喩) の解釈を扱う。属性比喩を扱った研究では、概念を属性値の集合で表し、顕現性について定量的計算を行う研究が主流である。岩山らは、プロトタイプ概念を表す属性値集合を確率的概念記述で記述する計算モデルを提案している [5]。Utsumi らや今井らは、認知心理学実験によって得た特徴の評定値を用いて属性値集合を表現している [6, 7] また、Martin による、実際の文章に現れる比喩表現を扱う研究 [8] や、梶井らによる、大規模なテキストを利用して自動構築した属性値集合を使って比喩解釈を目指す研究もある [4]。

属性比喩の解釈過程は、比喩表現において強調される特徴 (顕現属性値) が、たとえる概念 (source 概念) からたとえられる概念 (target 概念) へ重ね合わせられ、クローズアップされることによる、と説明できる。例えば、「砂のような雪」という直喩の解釈で「細かい」や「さらさらした」という特徴がイメージされるのは、source 概念「砂」の顕現属性値「細かい」が、target 概念「雪」で強調される属性値としてクローズアップされて解釈される。故に、source 概念の顕現属性値ではない「新しい」はクローズアップされないし、クローズアップされても理解不能となる。このように、比喩解釈にとって属性値のクローズアップは非常に重要である。概念間で正しく共有属性値がクローズアップされなければ、比喩解釈に悪影響を与える。

属性値のクローズアップ失敗の主要因として、概念を表す属性値集合 (のランキング) の歪みがある。例えば、今井らは、認知心理学実験の評定値を用いて、概念間でクローズアップされる顕現属性値の抽出を試みているが、評定値分布がばらつく場合の抽出精度が問題であると述べている [7, 2]。梶井ら [4] は、提案した比喩性検出手法の考察において、計算上比喩性検出には成功していても、人間の直観と異なる属性値がクローズアップされる問題を挙げている。

この問題は、特に、属性値集合を自動構成する場合に顕著となる。梶井らの手法によると、連体修飾関係を持つ語の共起確率が、属性値集合の生起確率分布とよく一致していることが前提となっている。このことは、さらに、次のように考察できる。連体修飾 (装定) の関係にある修飾語と名詞は、主述の関係に置換可能 [9] なので、連体修飾関係 (形容詞・形容動詞-名詞) による語の共起関係を、概念-属性値の関係とみなして抽出する手法自体は妥当といえる。一方、山梨によると、「連体修飾によって特徴づけられる装体表現は、一回事象的で、かつ対象指向的な機能を持つ表現 [10]」であり、必ずしも典型的特徴を表すものではない。よって、共起頻度に基づく尤度のみでは顕現属性値を保証されず、属性値ランキングに歪みが生じる可能性がある。

以上のことから、誤った比喩解釈を回避するためには、クローズアップ属性値の選択誤りを発見する必要がある。しかしながら、これらの作業を人手で実施するにはコストが高く、大規模な評価実験を実施することは困難である。

そこで本論文では、比喩解釈においてクローズアップされた属性値の適合性判定手法を提案する。我々は、WWW を知識源とし、対象概念とクローズアップ属性値による表現生成能力を検証することによって属性値の適合性判定を行う。適合性判定では、定型パターンを利用した知識獲得の枠組を利用する。定型パターンを利用して知識を獲得しようとする試みはこれまでも報告されており、その有効性はある程度確認されている [11, 12, 13]。知識源として WWW を用いる理由は、人間判断に匹敵する広い知識網羅性と低い分野依存性を備えた知識の活用には、新聞記事では不十分と判断したためである。

以下、2 で提案手法の基本的考え方を述べ、3 では提案手法の詳細について説明する。4 で提案手法の妥当性検証のための実験を行い、その結果を考察する。

2 基本的な考え方

本章では、提案手法の基本的な考え方を述べる。

本論文でいう「正しいクローズアップ属性値」は、source 概念の典型性を表す顕現属性値から選択される。よって、source 概念についての適合性判定としては、選択された属性値が、概念の顕現属性値であるかどうかを検証すればよい。source 概念とその顕現属性値を用いれば、「砂みたいに細かい」や「ハチのように忙しい」のように、比喩ま

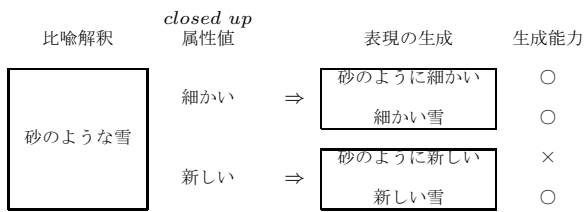


図 1: クローズアップ属性値と連体修飾と比較表現の関係

たは例示を意味する比較表現が生成できるはずである。ならば、source 概念と、クローズアップ属性値を用いて比較表現を生成し、その比較表現が一般に利用されているかどうかを調べれば、属性値の適合性を判定できる。このとき、比較表現を構成する「～のような」、「～のように」といった定型パターンを指標表現 (*ground*) と呼ぶ。

target 概念について属性値の適合性を考えると、比喩解釈においてクローズアップされる属性値は、概念の顕現属性値である必要はない。よって、target 概念における属性値の適合性については、概念の属性値として適しているかどうかを判定すればよい。また、連体修飾(装定)関係の機能を考慮すると、target 概念と属性値を組合せて典型的または一回事象的な連体修飾表現が生成できるはずである。よって、target 概念とクローズアップ属性値を用いて連体修飾表現を生成し、その連体修飾表現が一般に利用されているかどうかを調べれば、属性値の適合性を判定できる。

適合性判定には、比較表現の一般的な利用状況を把握できるだけの網羅性を持った知識源が必要となる。田添らの調査 [14] によると、新聞記事 1ヵ月分のコーパスから、指標表現「のような」を用いた比較表現が 500 組得られ、そのうち名詞句同士を比較したものは、311 組である。この場合も、10 年分の新聞記事から得られる比較表現は 37,000 組あまりにしかならない。指標表現「のように」についても同様の調査を行ったところ、10 年分の新聞記事から得られる比較表現は 20,000 組あまりに留まることがわかった。これらのことから、比較表現を用いて適合性判定を行うためには、新聞記事 10 年分程度では不十分であり、さらに大規模な知識源が必要であるという結論が得られる。

関口ら [15] によると、質に留意することによって、WWW データを新聞記事よりも網羅性が高く、かつ分野依存性が低い知識源として利用できる。本論文では、比較表現と連体修飾表現という単純な表現のみを扱うため、WWW を知識源を利用しても、質の低下等の影響を受け難いと考えられる。したがって、WWW を大規模でかつ、網羅性と低分野依存性を兼ね備えた知識源として利用できる。このような理由から、本論文では、適合性判定のための知識源として WWW を利用する。

3 提案手法

提案手法の概要を図 2 に、処理例を図 3 に示す。

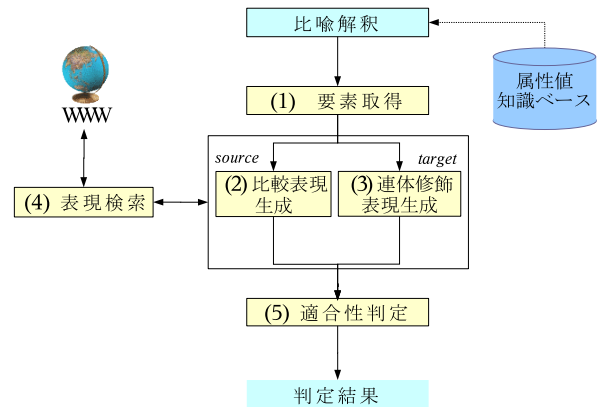


図 2: クローズアップ属性値の適合性判定手法の概要

適合性判定は、比喩性検出結果からの (1) 要素の取得、(2) 比較表現生成、(3) 連体修飾表現の生成、(4) 表現検索、(5) 適合性判定、からなる。

(1) では、比喩性検出結果から、source 概念、target 概念、クローズアップされた属性値を取り出す。(2) は、source 概念についての操作である。まず、「 $X_ground_γ$ 」というフレーム(定型パターン)を用意しておく。このとき、*ground* は指標表現であり、フレームの X には source 概念を、 $γ$ にはクローズアップ属性値を当てはめて比較表現を生成する。

指標表現 (*ground*) として、中村の調査 [16] による「比喩指標実現形」35 種類のうち、上で述べた比較表現を生成できるもの 9 種類¹ について WWW 中の頻度数を調べたところ、上位 2 組の指標表現「のように」、「のような」だけで全表現の頻度の 70% 以上を占めた。この結果から、本論文では、用意するフレームとして、「 $X_のような_γ$ 」、「 $X_のように_γ$ 」の二種類のみを用いる。

(3) は、target 概念についての操作である。クローズアップ属性値 $γ$ と target 概念 Y を「 $γ_Y$ 」というパターンで組合せ、連体修飾表現を生成する。(4) では、生成した比較表現、または連体修飾表現をクエリとして WWW 検索を行う。検索の結果得られた件数を、生成表現の頻度とみなして取得する。(5) では適合性判定を行う。表現検索で得られた頻度が閾値 $α$ 以上であれば、比較表現は生成可能であり、属性値は適合と判定される。頻度が $α$ 未満であれば、比較表現は生成不能であり、属性値は不適合と判定される。このとき、閾値 $α$ は、対象となっている属性値がクローズアップ属性値として適合か否かを判断するための基準である。 $α$ を大きくすると典型性を重視となり、小さくとると網羅性を重視となる。

図 3 では、「砂、雪」という概念対が処理対象であり、「砂」が source 概念、「雪」が target 概念である。適合性判定の

¹ 上位から順に、「のような」、「のように」、「みたいな」、「みたいに」、「に似て」、「に似た」、「のごとく」、「のごとき」、「じみた」。

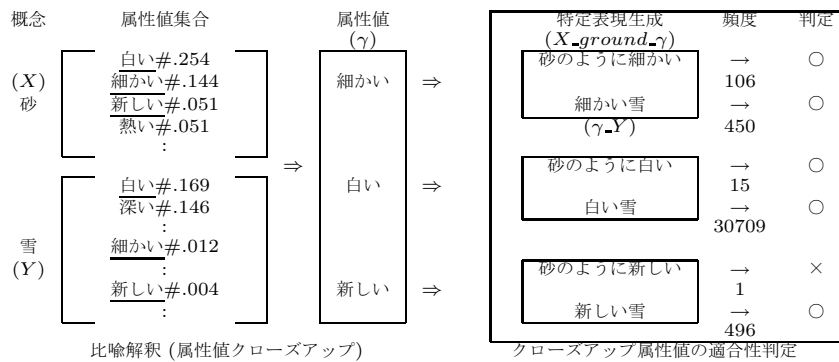


図 3: クローズアップ属性値の適合性判定の実行例

閾値は、 $\alpha = 20$ として考える。概念対に対する比喩解釈過程によって、属性値「細かい」、「白い」、「新しい」がクローズアップされている。

source 概念については、まず、フレーム「X_のような_γ」上の X に「砂」、γ に「細かい」「白い」「新しい」を適用して比較表現「砂のように細かい」、「砂のように白い」「砂のように新しい」を生成する。これらの比較表現を WWW 検索し、得た検索件数は、「砂のように細かい」が 106、「砂のように白い」が 15、「砂のように新しい」が 1 である。その結果、閾値を越える頻度が得られた属性値「細かい」、「白い」は適合、「新しい」は不適合と判定される。

target 概念については、まず、「γ_Y」の Y に「砂」、γ に「細かい」、「白い」、「新しい」を適用して「細かい雪」、「白い雪」、「新しい雪」という連体修飾表現を生成する。これらの連体修飾表現を WWW 検索し、得た検索件数は、「細かい雪」が 450、「白い雪」が 30709、「新しい雪」が 496 である。この結果、全ての属性値が適合と判定される。

4 実験と評価

適合性判定手法が、人間判断の代替と成り得るかどうかを調べるために、適合性判定能力を評価する。

まず、提案する属性値の適合性判定において、顕現属性値を決定する閾値 α の最適値を調べるための予備実験を行う。知識源として、新聞記事 10 年分² を用いて、基本知識ベースを構築した。ベースライン (BL) として、梶井ら [4] の比喩性検出手法を用いた。適合性判定における WWW 上のクエリ検索には、WWW 検索エンジンを利用した。

実験に用いる概念対として、梶井ら [4] の実験で用いた概念対 70 組を用いた。まず、概念対 70 組について、BL による比喩性検出処理を行い、クローズアップ属性値を得た。次に、得たクローズアップ属性値に対して、提案手法による適合性判定と、人手による適合性判定を行った。このとき、適合性判定の閾値 α を、 $\alpha = 1, 5, 10, 15, 20, 25, 30$ に変更し、7 種類の判定結果を得た。提案した適合性判定

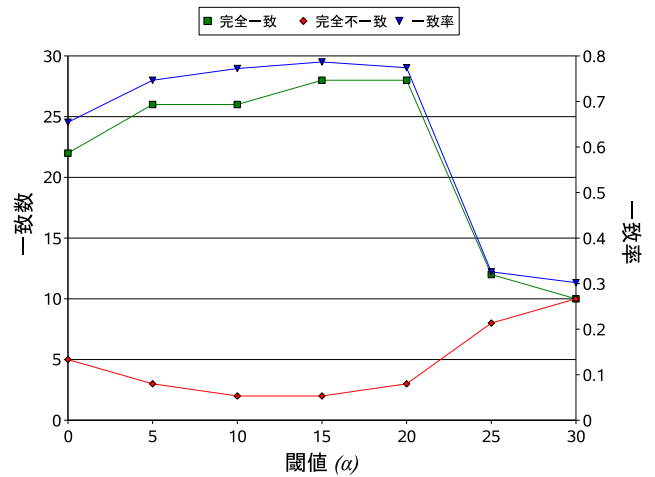


図 4: 適合性判定の閾値の変化とクローズアップ属性値の適合性判定の性能

(以下、適合性判定)の結果と、人手による適合性判定(以下、人手判定)の結果を比較し、一致率を求めた。

実験の結果を図 4 に示す。図には、両判定結果の一致率と適合性判定の閾値 α の関係 (▼)、適合性判定結果が人手判定結果と完全に一致した数 (□)、全く一致しなかった数 (◇) が示されている。

両手判定結果の一致率の推移を見ると、閾値 $\alpha = 15$ のときに、0.79 と、非常に高い値が得られた。また、 $\alpha \leq 20$ においては、0.65~0.79 と、概ね良い一致を見せており、本手法が、有効であることを示している。一方、 $\alpha \geq 25$ になると一致率は 0.35 未満と急激に低下した。

5 考察

適合性判定結果と人手判定結果の一致率が $\alpha \geq 25$ で急激に低下する理由について考察する。一致率の低下傾向は、適合性判定結果と人手判定結果が完全に一致した数の推移と非常に類似しており、 $\alpha \geq 25$ の条件で急激に減少

² 毎日新聞 CD-ROM 版 1994 年~2003 年、約 111 万記事 (約 1046 万文) を用いた。

表 1: クローズアップ属性値の適合性判定結果の例

概念対	クローズアップ属性値	適合性判定 ($\alpha = 15$)	人手判定	一致率
人形, 表情	美しい	○	○	1.000
	可愛い	○	○	
	かわいい	○	○	
	いい	×	×	
	不思議な	×	×	
	ユーモラスな	×	×	
子供, 心	素朴な	×	×	.500
	グロテスクな	×	×	
	新しい	×	×	
	個性豊かな	×	×	
	いい	○	×	
	やすい	×	×	
夢, プロジェクト	悪い	×	×	.250
	健康な	×	○	
	高い	○	×	
	弱い	○	×	
	不自由な	×	×	
	幼い	○	×	

している。しかし、全く一致しなかった数の推移はこれと相反する傾向を示している。このことから、 $\alpha \geq 25$ の条件では、過剰判定によって有効な属性値が不適合と評価され、その結果一致率が低下したと考えられる。

表 1 に、適合性判定結果 ($\alpha = 15$) と人手判定結果の例を示す。左から、対象概念対、BL によってクローズアップされた属性値、適合性判定結果、人手判定結果、両判定結果の一致率を示している。「人形, 表上」の概念対に対しては、「美しい」「可愛い」という典型的特徴を示す属性値が適合と判断され、「不思議な」「グロテスクな」「新しい」などの一回記述的な特徴となり得る属性値は不適合と判断されており、人間の判断に近い結果が得られている。

一方、人手判断と一致しなかった結果のほとんどは、以下のようなケースであった。「神様, 人」では、「偉い」という属性値がクローズアップされており、人手判定では適合とされている。適合性判定でも、WWW で実際に「神様のように偉い」と記述されているケースがあるにもかかわらず、取得した頻度が 7 であるため、閾値によって不適合となってしまった。また、「枝, 柱」では、「太い」という属性値がクローズアップされており、人手判定では不適合であるが、適合性判定では適合となっている。これは、比較表現「枝のように太い」の WWW 検索において、「…、木の枝のように太い根から細い根が枝分かれています。」や、「緑化樹剪定枝のように太い枝が入っていたり、…」のように、意味の異なる表現の一部にも一致してしまい、取得される頻度が実際より大きな値になることが主な原因であると考えられる。これらの問題については、適合性判定の基準として静的な閾値によって一元的に区別してしまう点に工夫の余地がある。検索された文を統語構造解析し、統語範囲を把握するなどの対応が考えられる。

一致率が 0.5 を下回ったケースが 4 組あったが、全て「夢」を source 概念とする概念対であった。これは、田添らの分析 [14] にあるように、「夢」という概念そのものが

比喩を生成する性質を持つため、顕現属性値を明示的に示した比較表現が記述されにくいためと考えられる。

6 おわりに

本論文では、比喩解釈において強調される特徴として、概念間でクローズアップされる属性値の適合性を自動判定する手法を提案した。実験による検証の結果、自動評価では、約 80% 人間の判断をシミュレートでき、人間判断の代替としての利用も考慮可能であることがわかった。

今後は、提案した適合性判定の結果を、属性値集合へ反映する手法について検討する予定である。

参考文献

- [1] K. Shinohara and B. Indurkha. Braiding the Multiple Threads of Interdisciplinary Research on Metaphor. 日本認知科学会第 17 会大会発表論文集, pp. 292–297, 2000.
- [2] 今井豊, 石崎俊. 比喩理解のための顕現属性の幾何学的抽出法. 人工知能学会誌, Vol. 15, No. 2, pp. 309–320, 2000.
- [3] 内山将夫, 村田真樹, 馬青, 小作浩美. 統計的手法による換喩の理解. 自然言語処理, Vol. 7, No. 2, pp. 91–116, 2000.
- [4] 樹井文人, 福本淳一, 椎野努, 河合敦夫. 確率的尺度を用いた比喩性判定手法. 自然言語処理, Vol. 9, No. 5, pp. 71–92, 2002.
- [5] 岩山真, 徳永健伸, 田中穂積. 比喩を含む言語理解における顕現性の役割. 人工知能学会誌, Vol. 6, No. 5, 1991.
- [6] Akira Utsumi, Koichi Hori, and Setsuo Ohsuga. An Affective-Similarity-Based Method for Comprehending Attributional Metaphors. *Journal of Natural Language Processing*, Vol. 5, No. 3, pp. 3–31, 1998.
- [7] 今井豊, 石崎俊. 比喩理解における顕著な属性の発見手法. 自然言語処理, Vol. 6, No. 5, pp. 27–42, 1999.
- [8] J. H. Martin. Metabank: A Knowledge-Base of Metaphoric Language Conventions. *Computational Intelligence*, Vol. 10, No. 2, pp. 134–149, 1994.
- [9] 佐久間鼎. 修飾の機能. 表現文法 (日本文法講座 5), pp. 23–55. 明治書院, 1958.
- [10] 山梨正明. 認知文法論. 言語の経験値基盤とイメージスキーマ, 第 4 巻. ひつじ書房, 1995.
- [11] Marti A. Hearst. Automated Acquisition of Hyponyms from Large Text Corpora. In *14th International Conference on Computational Linguistics (COLING92)*, 1998.
- [12] Matthew Berland and Eugene Charniak. Finding Parts in Very Large Corpora. In *International Conference on Association of Computer Linguistics (ACL-1999)*. ACL, 1998.
- [13] 安藤まや, 関根聡, 石崎俊. 定型表現を利用した新聞記事からの下位概念単語の自動抽出. 情処研報 NL157-11, pp. 77–157, 2003.
- [14] 田添文博, 椎野努, 樹井文人, 河合敦夫. “名詞 a のような名詞 b” 表現の比喩性判定モデル. 自然言語処理, Vol. 10, No. 2, pp. 43–58, 2003.
- [15] 関口洋一, 山本和英. Web コーパスの提案. 情処研報.
- [16] 中村明. 比喩表現の理論と分類. 共立出版, 1977.